Biochimica et Biophysica Acta 1817 (2012) 1256–1262

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbabio

Empirical and computational design of iron-sulfur cluster proteins ☆

Joanna Grzyb [a,*], Fei Xu [b], Vikas Nanda [b], Renata Łuczkowska [a], Eduard Reijerse [c], Wolfgang Lubitz [c], Dror Noy [d]

[a] Institute of Physics, PAS, al. Lotników 32/46, 02–668 Warsaw, Poland
[b] Biochemistry Department, Robert Wood Johnson Medical School, Univ. of Medicine and Dentistry of New Jersey, Piscataway, NJ, USA
[c] Max-Planck-Institut für Bioanorganische Chemie, Mülheim an der Ruhr, Germany
[d] Weizmann Institute of Sciences, Rehovot, Israel

ARTICLE INFO

Article history:
Received 16 November 2011
Received in revised form 13 January 2012
Accepted 1 February 2012
Available online 8 February 2012

Keywords:
Iron-sulfur cluster
Heme
Protein design
Empirical design
Computational design
Maquette

ABSTRACT

Here, we compare two approaches of protein design. A computational approach was used in the design of the coiled-coil iron-sulfur protein, CCIS, as a four helix bundle binding an iron-sulfur cluster within its hydrophobic core. An empirical approach was used for designing the redox-chain maquette, RCM as a four-helix bundle assembling iron-sulfur clusters within loops and one heme in the middle of its hydrophobic core. We demonstrate that both ways of design yielded the desired proteins in terms of secondary structure and cofactors assembly. Both approaches, however, still have much to improve in predicting conformational changes in the presence of bound cofactors, controlling oligomerization tendency and stabilizing the bound iron-sulfur clusters in the reduced state. Lessons from both ways of design and future directions of development are discussed. This article is part of a Special Issue entitled: Photosynthesis Research for Sustainability: from Natural to Artificial.

1. Introduction

The growing human population increases global energy demands, while fossil fuel supplies are rapidly depleting. It is now commonly accepted, that improvement of photosynthesis yield or even bioengineering, and rearranging known photosynthetic pathways may enable the use and storage of solar energy in more efficient ways. Artificial biomimetic and bioinspired devices may be used to provide alternative energy sources. Understanding and, to a greater extent, engineering of photosynthesis, calls for new types of proteins that may be provided by *de novo* design methods.

Empirical design is the earliest method, and was first applied to create and optimize heme binding proteins [1]. This led to the so-called protein maquette: a well defined, robust protein, with specific functionality. As protein stability is a factor of great importance, the α-helix was the structure of choice. Amino acid sequences approximated by heptad repeats combined with binary patterning of polar and nonpolar amino acids, resulted in water soluble helical bundles [2]. Functionalization of the maquette may be achieved in several ways; one of the widely studied is heme binding by two histidine residues in the hydrophobic core of the bundle. Detailed research enabled finding several factors influencing the affinity, heme redox properties, and binding stoichiometry. The same proteins or their derivatives bind also other porphyrins and chlorins [3]. Several designed proteins served as scaffold for metal binding, e.g. single or di-atomic binding sites for iron [4,5], zinc [6] or cobalt [7,8] as well as more complicated iron-sulfur clusters [9–11].

The second approach is computational design that, in general, relies on several algorithms optimizing structure for minimal global energy [12]. Using the computational power of computer clusters, one can generate libraries of structures within selected constraints and estimate their potential usefulness [13]. Some constraints can be formulated *a priori*, such as the introduction of [4Fe-4S] binding site in CCIS [14]. However, more random simulations may scan a wider probability space. In such a case, structures may be evaluated *a posteriori*, e.g., for presence of specific enzymatic activity [15].

The computational method enables designing new binding sites within existing proteins [4], protein-protein interfaces [16], and non natural folds [17], modeling receptors for novel ligands [18], and creating enzymes within new scaffolds [15,19], catalyzing reactions that do not exist in nature [20,21]. Computational design is especially powerful in combination with directed evolution [22], when new functionality can be easily tested, and selected variants subjected to a second round of design and optimization.

Not withstanding the successful examples listed above, protein *de novo* design has many challenges. The first on the list is the prediction

Abbreviations: CCIS, coiled-coil iron-sulfur protein; DTNB, 5,5'-dithiobis-(2-nitrobenzoic acid); RCM, redox-chain maquette.

☆ This article is part of a Special Issue entitled: Photosynthesis Research for Sustainability: from Natural to Artificial.

* Corresponding author. Tel: +48 505147798; fax: +48 22 843 09 26.
E-mail address: jgrzyb@ifpan.edu.pl (J. Grzyb).

0005-2728/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbabio.2012.02.001

of protein dynamics and structural changes due to ligand binding. As a result, new designs often do not perform exactly as intended. Designing β structures is still more complicated than that of α-helices. Another serious challenge is accounting for electrostatic interactions, which may strongly determine structure stability, and may lead to oligomerisation or domain swapping. Designing loops, which partially determine structure stability, but may contain functional elements, is yet another difficult challenge [23].

Here we compare the design of two iron-sulfur cluster proteins; one using a computational strategy, and the other based on empirical knowledge. These types of proteins are of special interest in understanding photosynthesis [24] and mitochondrial oxidation [25], as well as in biotechnological processes, such as bio-hydrogen production by hydrogenases [26,27]. These proteins may also be used to improve or control photosynthesis *in vivo* or as a part of stand-alone systems. The coiled-coil iron sulfur cluster protein (CCIS) was designed by a computational method as a novel fold without natural analogs. This protein is a four-helix bundle that binds a four-iron four-sulfur [4Fe-4S] cluster in the middle of its hydrophobic core, while in natural proteins iron-sulfur clusters are coordinated by residues (mostly cysteines) generally from β-sheets and loops. The redox chain maquette (RCM) design was based on a wide range of experimental data, and lessons learned from previously published protein designs. The RCM fold is also a single-chain four-helix bundle, but with a hydrophobic core for binding one porphyrin ligand, and iron-sulfur cluster binding sites in connecting loops. The first CCIS prototype, CCIS1, was reported previously [14]. Here, we examine CCIS1 and several new variants of it designed in order to improve iron-sulfur cluster stability, as well as the first two prototypes of RCM. The study of their assembly with cofactors and its effect on the electronic properties of bound cofactors, the lessons in designing functional, cofactor-binding proteins, and their implementation in the next generation of CCIS and RCM proteins are discussed here.

2. Materials and methods

2.1. Empirical design of Redox-Chain Maquettes (RCM)

The design of the RCM family aimed at three protein structural features:

1. A single chain bundle of four amphiphilic helices
2. A binding site for heme or other porphyrin derivatives in the middle of the bundle's hydrophobic interior
3. Assembly of [4Fe-4S] clusters in the loops connecting the helical domains

In order to confer these structural features on a protein sequence, we followed a set of empirical rules. Helix propensities of amino acids, including different variants of sequence and residues neighborhoods (see e.g. [28,29]) enable reasonable predictions of helix stabilities. The behavior of amphiphilic helices in aqueous solution is also predictable and determined by a binary pattern of hydrophobic and hydrophilic residues [2]. Additionally, charged residues were placed in patches on opposite helices in order to stabilize their desired orientation. Heme or porphyrin binding is introduced by placing of two histidines in two parallel helices facing the bundle's hydrophobic core and providing axial ligands to the heme/prophyrin central metal (Fig. 1a). Following Ghirlanda et al. [30] small residues (glycines) were placed on the respective side of the two helices that do not carry histidines in order to provide enough space for the porphyrin ring. For iron-sulfur cluster binding, two types of loop topologies were introduced: An asymmetric loop containing the CxxCxxCxxxxC motif that is based on the [4Fe-4S] binding site from bacterial ferredoxin [31,11], and a symmetric two-loop topology inspired by the Fx binding site of photosystem I [32] whereby each loop contains half the binding site in a CxxxxxC motif (Fig. 1a).

The electron transport chain of formate reductase, a huge protein complex containing both heme and [4Fe-4S] clusters inspired the cofactor assembly in RCM. In order to keep the cofactors within effective electron tunneling range , the length of the helices was set up to be not more than 25 Å, which implies an edge-to-edge distance of 6–8 Å between heme and FeS clusters. Additionally, sequence analysis of formate dehydrogenase from different species revealed a highly conserved aromatic residue, which orients the heme plane toward the FeS cluster. For this purpose, we introduced tryptophans at position 16 and 80. The distinct near-UV absorption band of tryptophan is also a very useful marker for protein tracking during purification. The final design (Fig. 1a) was 101 amino acids long with a molecular weight close to 12 kDa.

Structural modeling using I-tasser (an automated platform, using multiple-threading alignments [33]), as well as preliminary experimental data suggested that the long asymmetric loop may fold into a short α-helix, which may prevent [4Fe-4S] assembly, and lead to high oligomerization tendency and low affinity to heme. In order to test this, we created two RCM variants:

I. RCM1, with Cys residues substituted with Ser within the asymmetric loop
II. RCM2, with substitution of the whole asymmetric loop by a strictly hydrophylic, flexible GGSGSGG sequence.

2.2. Computational design of Coiled-coil iron-sulfur cluster proteins (CCIS)

The first CCIS prototype, CCIS1 was recently designed using the ProtCad software [34] using a metal-first computational approach [14]. The design resulted in a completely new fold, without any known natural analogs, in which the [4Fe-4S] cluster is bound within the hydrophobic core of an all α-helical coiled coil. Several variants were designed in order to improve the properties of the first prototype (Fig. 1b)

a) CCIS1a was designed to increase the stability of the reduced state of the [4Fe-4S] cluster by improved packing of the bundle. In comparison to the original maquette, the buried glutamate (E36) was substituted by serine, and the tryptophan at the helix interface (W14) was substituted by alanine. Additionally, the tryptophan at position 96, the possible place for aggregation, was substituted by serine. These three mutations were present in all following variants, with the exception of lysine/glutamate instead of serine at position 96 of CCIS2c/2-d, which provided amino acids of opposite charges (E-K and E-R pairs) in the 8th and 22nd position of each helix.
b) CCIS2b carried five mutations over the helices, which should improve their helix propensity.
CCIS2c and 2-d addressed the problem of preventing oligomerization, by changing the net charge of the protein surface.
c) CCIS2c was achieved by substitution of eight selected residues at the hydrophylic surface of the bundle to basic K or R.
d) CCIS2d addressed charge pairing optimization of the bundle and was achieved mainly by substitution of the same residues to acidic E.

2.3. Protein expression and purification

Genes of RCM and CCIS1 were synthesized by Bio S&T(Montreal, Canada) and cloned into pet32b (Novagen). Mutations were introduced by an insertion-deletion protocol (Finzymes) and confirmed by sequencing. Expression and purification were done as described in [14]. The only exception was the use of HiTrapS (GE Healthcare) instead of HiTrap Q columns in the ion exchange step of RCM1 and CCIS2c purification.

Finally, all proteins were desalted to 25 mM Tris/HCl pH 8.0 and stored in aliquots at -20 °C.

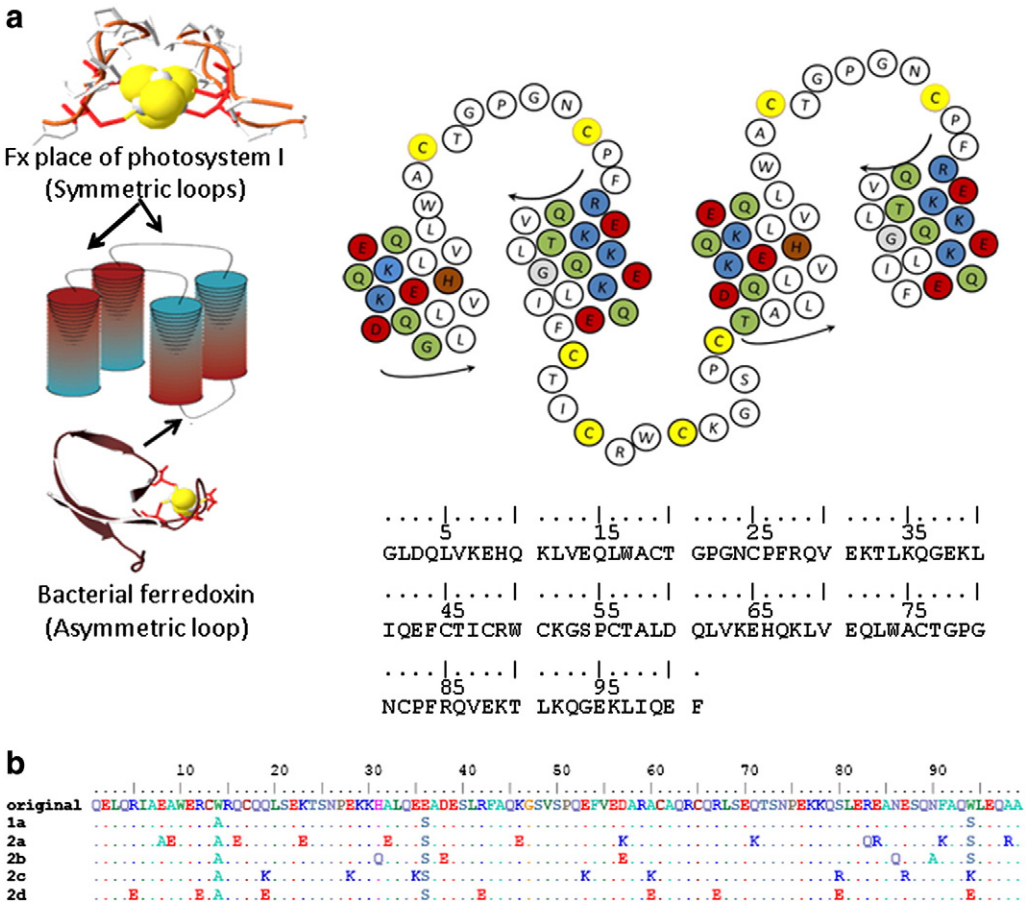

Fig. 1. The templates and results of design: (a) RCM: scheme of desired native RCM structure, Fx place of photosystem I (note symmetry and positions of red-marked cysteinyl residues), bacterial ferredoxin loop with [4Fe-4S] cluster and final sequence, together with helical diagram representation (helices colored as follows: brown-histidine, light gray–glycine residues placed to provide enough space for the heme ring, red-acidic, blue-basic, green-polar); (b) Alignment of previously characterized CCIS1 (original) and its variants, as described in the text. Dots represent residues identical to the template.

2.4. Reconstitution of iron-sulfur clusters

Two protocols were used for iron sulfur assembly, clusters both of which were carried out under strict anaerobic conditions. The short assembly protocol [14] was performed at room temperature. Degassed proteins were first incubated for about 30 min under anaerobic atmosphere and to assure complete deoxygenation. Protein concentrations of 50, 200, 1000 μM, were tested. The assembly buffer was 50 mM Hepes/NaOH, pH 8.0. In addition, Hepes/NaOH, pH 8.5, Hepes/NaOH + 0.5 M NaCl, pH 8.0, and Tris/HCl pH 8.3 with or without 0.5 M NaCl were tested. DTT was added to the solution, at a molar ratio of 2.1:1 (DTT:molar concentration of cysteine residues). After 30 min, $FeCl_3$ and Na_2S were added to the mixture (1.2:1 molar ratio, $FeCl_3$:cysteine). After 45 min of incubation with gentle stirring, the mixture was passed through a PD10 desalting column (GE Healthcare), and the colored fractions (red-brown) were collected and examined. Prepacked columns were kept under anaerobic conditions and equilibrated with the same buffer for the assembly. Alternatively, columns were packed anaerobically using the Sephadex G-25 powder and degassed buffer. The long protocol [35] was carried out at 4 °C, overnight (16-18 h). Protein solutions (10–40 μM) in buffer (50 mM Tris/HCl pH 8.3) with added β-mercaptoethanol (4% v/v) were degassed and incubated with stirring for 30 min, followed by addition of ferric nitrate and Na_2S. After overnight incubation, the solution was concentrated 20 fold using a stirred cell with a 1 kDa cut-off membrane (Millipore), and passed through a PD10 desalting column that was equilibrated with a 50 mM Tris/HCl pH 8.3 buffer with 15% glycerol. Red-brown fractions were collected and analysed.

2.5. Heme binding

All RCMs variants were tested for heme binding. Heme solutions (usually about 1 mM) were prepared in 0.1 mM NaOH and diluted prior to use to concentrations adequate to protein titration. The concentration of heme was determined as described in [36]. The titrations were done manually with a microsyringe. UV-Vis Spectra were recorded with a Cary50 Bio spectrophotometer.

2.6. Size exclusion chromatography

Proteins were analysed by size exclusion chromatography with a Superdex 75 5/150 column (GE Healthcare) equilibrated with 50 mM Tris/HCl pH 8.0 + 0.5 M NaCl. To assure anoxygenic conditions, the buffer was degassed and kept under a stream of nitrogen during the run, and the column was equilibrated with 3 column volumes of degassed buffer before loading a sample. The elution was monitored at three wavelengths: 280 nm (protein), 350 nm (iron-sulfur cluster absorption band, iron bound to protein) and 415 nm (exclusive indication of iron-sulfur cluster). In heme containing preparations, the 415 nm absorption is also indicative of bound heme and the 350 nm is a nonspecific absorption of heme, both bound and unbound to protein.

2.7. Circular dichroism

Circular dichroism spectra in the near UV or visible range were recorded with Jasco J-815 spectropolarimeter. For UV (175–259 nm)

range measurements, proteins were desalted to 10 mM phosphate buffer, pH 8.0 and cuvettes of 0.1 or 1 mm optical pathlength were used. For visible range measurements, 4 mm or 10 mm optical pathlengths were applied. Contents of secondary structure elements were calculated with the CDPRO software [37], based on a library of water-soluble proteins.

2.8. Electron paramagnetic resonance (EPR)

Low temperature X band EPR spectra were measured on a Bruker ELEXSYS 500 spectrometer in 3.4 mm (i.d.) capillary tubes. Temperature was controlled by a ESR900 cryostat cooled by liquid helium (Oxford Instruments). EPR signals from 2,2-diphenyl-1-picrylhydrazyl (DPPH) were used as a reference for determining *g* values

2.9. Chemical modification of thiol groups

Thiol groups of the protein were blocked by reversible modification with dinitrobenzoic acid (DTNB, Sigma) [38]. Protein solutions in 50 mM Tris/HCl pH 8.0 were incubated with 10 times molar excess of dithiotreitol for 15 min in RT, with gentle mixing. DTNB (from a 200 mM stock in the same buffer) was added to reach 30 times molar excess (compared to the protein) and incubated for 15 min. After incubation, the modified proteins were purified on a HiTrap desalting column (GE Healthcare) equilibrated with the same buffer. To confirm the successful modification, aliquots of modified proteins were reduced again with DTT, and the amount of released yellow ion of TNB was determined spectrophotometrically.

3. Results

3.1. Secondary structure

The UV-CD spectra of apo- and holo-proteins are shown in Fig. 2. The spectra of RCM1 (not shown), and RCM2 (Fig. 2a) apoproteins were almost identical, typical of α-helical proteins, and very similar to other known four-helix bundle proteins [30]. The spectra remained essentially unchanged following heme incorporation, and assembly with [4Fe-4S] clusters. Blockage of thiol groups, which prevents oligomerization by disulfide bridge formation, only slightly increased the random-coil features (not shown) of the spectrum. Similarly, the CD spectrum of apo-CCIS2b was typically α-helical with little or no change upon assembly with the [4Fe-4S] cluster (Fig. 2c). The spectra of CCIS2a, CCIS2c and CCIS2d (not shown) were similar. In contrast, the CD spectrum of the CCIS1 apoprotein was indicative of partially unfolded protein, but after [4Fe-4S] assembly, the α-helical character of the spectrum increased significantly (Fig. 2b). In comparison to CCIS1, the CD spectrum of CCIS1a indicated a large amount of unstructured peptide (not shown).

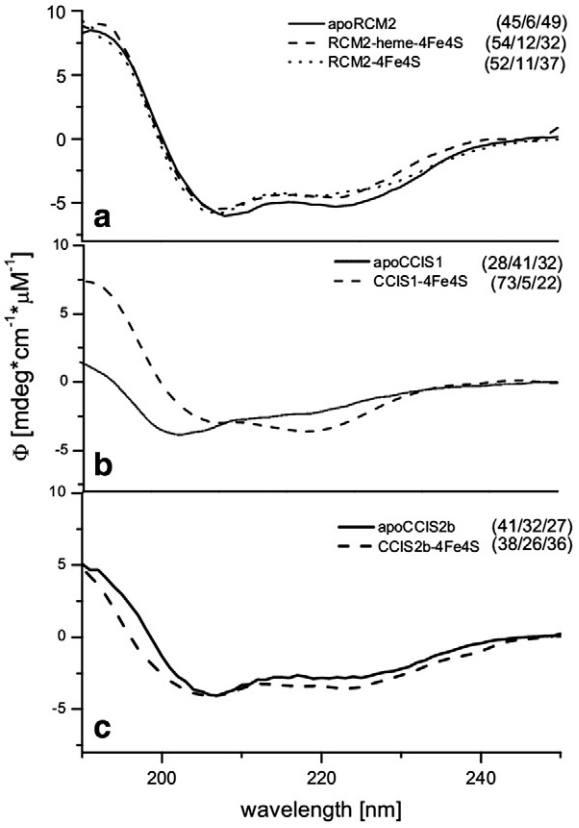

Fig. 2. UV circular dichroism spectra of (a) RCM2 apoprotein, RCM2 assembled with only with [4Fe-4S] or [4Fe-4S] and heme, (b) CCIS1 apoprotein and CCIS1 with [4Fe-4S] cluster, and (c) apoprotein of CCIS2b and CCIS2b with [4Fe-4S] cluster. Measurements on apoproteins were done for a 50 μM protein concentration in 10 mM phosphate buffer, pH 8.0, using 0.1 mm optical pathlength; assembled proteins were measured in a gas tight cuvette, 1 mm optical pathlength. Numbers in brackets represent calculated percentage of secondary structure elements, helix/sheet + turn/unordered.

3.2. Cofactors assembly

Our previous study [14] demonstrated the incorporation of a [4Fe-4S] cluster into the hydrophobic interior of CCIS1. Similarly, the holoprotein UV–VIS spectra of all newly designed CCIS variants featured the typical iron-sulfur cluster signature – a broad band at 400–415 nm, and a shoulder at about 350 nm (Fig. 3). The ratio of absorbance at 415 nm to 280 nm is indicative of the yield of assembly. It varied slightly between repetitions, but usually was about 0.5 in CCIS2b, 0.35 in CCIS2a, CCIS2c, and CCIS2d, and lower than 0.2 in CCIS1a. The absorbance ratio in CCIS1 was 0.2 but CCIS1 has three tryptophan residues whereas the other variants have only one. Thus, the assembly yields of CCIS1 and CCIS2b are comparable, whereas CCIS2a, CCIS2c and CCIS2d have slightly lower yields, and CCIS1a has the lowest yield. The assembly yields were similar in both the long and short reconstitution protocols.

In RCM1 and RCM2, the short reconstitution protocol was efficient, yielding the same typical [4Fe-4S] absorption spectrum for both proteins (Fig. 4a). The 415 to 280 nm absorbance ratio was close to 0.5, and the extinction coefficient at 415 nm was found to be 21750 M^{-1} cm^{-1}, which is in the range expected for natural [4Fe-4S] cluster proteins (16000–23000 M^{-1} cm^{-1} [39]). Titration of these holoproteins with heme resulted in the expected binding curve reflecting a 1:1 heme:protein stoichiometry. The heme dissociation constant (Kd) was found to be 10.3 ± 0.16 μM by fitting the curve with an

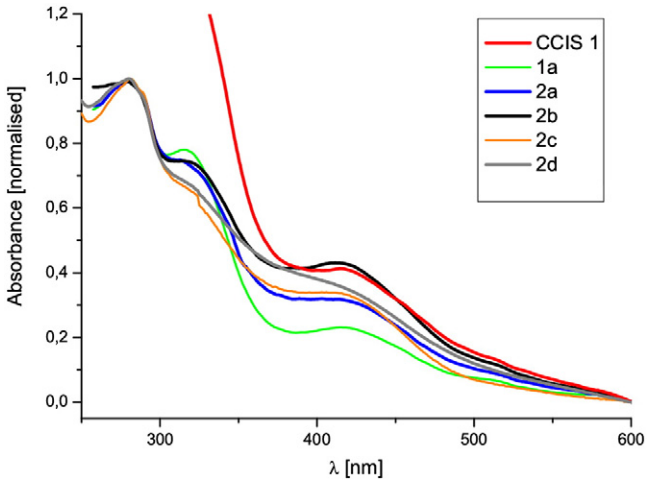

Fig. 3. The absorption spectra of assembled CCIS protein variants, normalized to the number of tryptophans per peptide chain (3 for CCIS1 and 1 for the other variants). Spectra were measured right after assembly in gas-tight cuvettes.

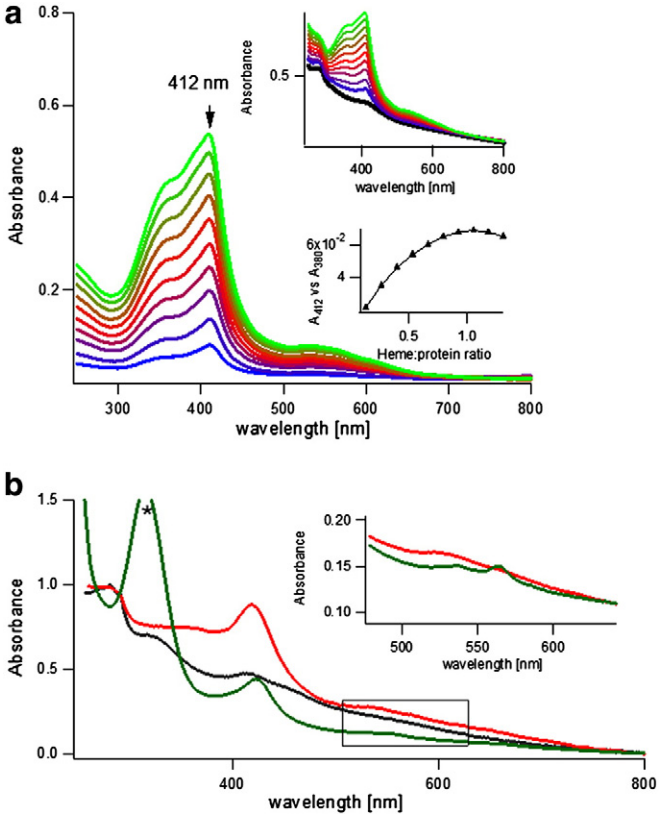

Fig. 4. Titration of heme into RCM2-[4Fe-4S] (a) and assembly of [4Fe-4S] into RCM2 with previously bound heme (b). Panel (a) presents UV–VIS absorption spectra of several titration steps using the heme-free RCM2-[4Fe-4S] spectrum for baseline correction; insets – uncorrected spectra and the 412 nm vs 380 nm absorption difference trace. Panel (b) presents UV–VIS absorption spectra of RCM2 assembled with [4Fe-4S] cluster only (black), and with both [4Fe-4S] cluster and heme (red). Changes in FeS- heme reconstitution after ditionite addition are in green. The asterisk marks the absorption of dithionite, inset – zoom into the heme Q band region at 450–650 nm.

equilibrium equation for a one step-binding model. Subtracting the initial [4Fe-4S] cluster protein spectrum from the spectra of the proceeding titration steps yielded typical cytochrome b-type heme-protein absorption spectra with an absorption maximum at 414 nm (Fig. 4a). This implies that heme binding does not affect the bound iron-sulfur cluster. Furthermore, it was possible to follow heme binding with iron-sulfur cluster reconstitution into RCM proteins. The heme absorption peak at 414 nm was unchanged, and overlapped with the [4Fe-4S] spectrum (Fig. 4b). Reduction of the sample with sodium dithionite decreased the [4Fe-4S] absorbance and shifted the heme absorbance to 424 nm, and induced a characteristic change in the heme Q absorption band around 550 nm. These features are typical of heme and [4Fe-4S] cluster reduction.

Interestingly, when either the RCM1 or RCM2 apoproteins were titrated the absorption peak maximum bound heme started at 424 nm in the first few titration steps, and gradually shifted to 414 nm with the addition of heme (see supplementary, Fig. S1). This was not observed in proteins that had their SH thiol groups blocked with DTNB. In the unmodified proteins, the stoichiometry of binding was close to 1:1 (heme:protein, Fig. 4), and the heme Kd was found to be 2.4±0.8 μM, and 0.36±0.2 μM for RCM1, and RCM2, respectively. These values are quite high and represent low affinity, when compared to other heme binders. However, we found the heme binding affinity to significantly increase in DTNB modified proteins, up to 0.14±0.03 μM, and 0.23±0.02 μM for RCM1, and RCM2, respectively. The high standard deviations are most likely related to oligomerisation tendencies, discussed below.

The EPR spectra of all CCIS variants were indicative of the presence of a [4Fe-4S] cluster (rhombic shape, characterized by g values of 1.90, 1.97 and 2.04). However, the spectra show low intensity and the signal intensity did not increase linearly with concentration. In some cases of concentrated samples (0.8-1.2 mM), degradation products of the cluster were evident in the EPR measurements. Addition or removal of glycerol did not improve the signal quality. For RCM2 assembled with an iron-sulfur cluster, the EPR spectrum showed a [4Fe-4S] signature (not shown). Surprisingly, no signature of heme was found for the RCM2-[4Fe-4S]-heme assembly. This suggests that the heme was coupled somehow to the [4Fe-4S] cluster to become EPR silent. The secondary structure was not affected for any of the holoprotein samples (not shown).

3.3. Oligomerization states

Size exclusion chromatography of apo-RCM1 revealed a mixture of species dominated by dimers and higher mass oligomers even under disulfide reducing conditions (supplementary Fig. S2). The minor fraction of apo-RCM1 monomers eluted at a volume corresponding to 18 kDa. In contrast, apo-RCM2 was about 70% monomeric, and converted to 100% monomers after reducing the disulfide bridges with DTT. The monomer fraction eluted at a volume corresponding to 22 kDa. The elution volumes of both RCM1 and RCM2 monomers were higher than expected for globular proteins with the same molecular weights suggesting that RCM1 and RCM2 had a rod-like structure.

RCM2 reconstituted with [4Fe-4S] was strictly monomeric (Fig. 5a), whereas heme binding without [4Fe-4S] incorporation resulted in purely dimeric proteins (Fig. 5b). In the elution profile of RCM2 reconstituted with both heme and [4Fe-4S] clusters, both protein dimers and monomers were present (Fig. 5c). RCM1 reconstituted with either [4Fe-4S] or heme was a mixture of dimers and higher oligomers (not shown).

The apoproteins of all CCIS variants except CCIS2d were found to be monomeric but eluted differently on the size exclusion column (supplementary Fig. S3). While CCIS1 was eluted, as previously described, with an elution volume corresponding to about 18 kDa, CCIS1a, CCIS2a, CCIS 2b, and CCIS2c had retention volumes corresponding to 24–27 kDa. These differences imply that apo-CCIS1 has a more globular shape than the other variants. CCIS2d eluted at a volume corresponding to 36 kDa indicating that the protein formed a very stable dimer.

The oligomerization state of CCIS holoproteins varied significantly according to the reconstitution protocols. Samples assembled according to the *short protocol* (high protein concentration, assembly at room temperature for 1 h) were mostly dimeric or oligomeric, whereas samples assembled according to the *long protocol* (low protein concentration, assembly at 4 °C for 16-24 h) were almost exclusively monomeric (Fig. 5d). The elution volume of monomeric holoproteins was exactly the same as the respective apoproteins, suggesting that assembly did not significantly change the shape of the protein.

4. Discusssion

The CCIS and RCM proteins are representative outcomes of two complementary approaches of protein *de novo* design that were used for addressing the highly complicated problem of designing cofactor-binding proteins. The empirical approach that was used for RCM design is more intuitive and relies on combining natural binding motifs or their fragments. In RCM, native-like FeS cluster binding loops were grafted on a heme binding four-helix bundle motif that was originally inspired by natural cytochrome b. The computational approach is not restricted to natural motifs and should be capable of designing non-natural folds as in the case of CCIS-type proteins. The designs presented here demonstrate the advantages and disadvantages of each approach.

Both methods were fairly successful in designing global protein folds. All CCIS and RCM variants were α-helical as designed. In

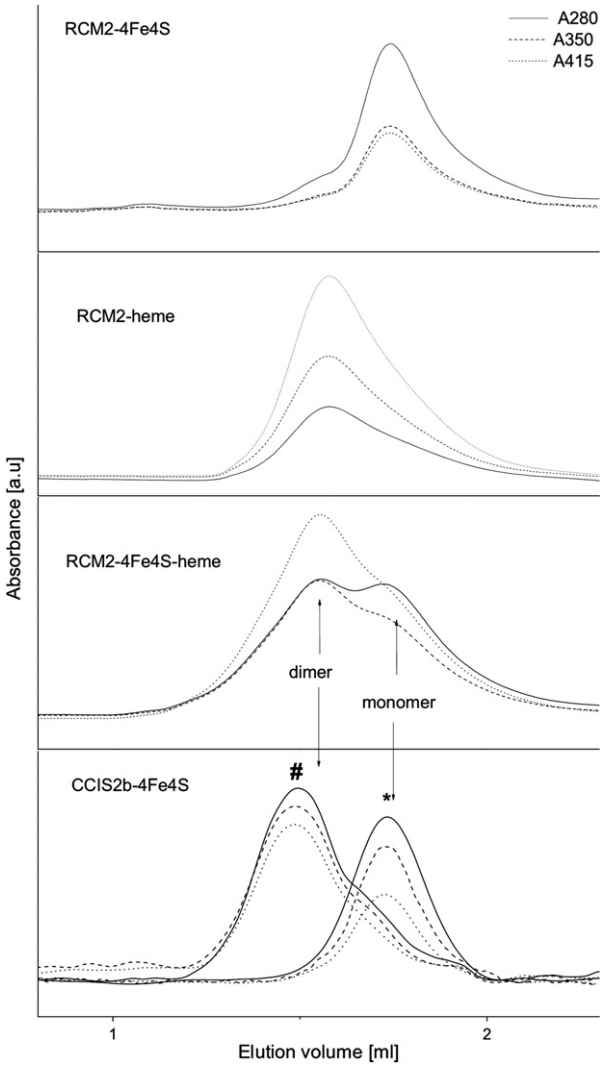

Fig. 5. Elution spectra of RCM2-[4Fe-4S], RCM2-heme, RCM2-[4Fe-4S]-heme and CCIS2b-[4Fe-4S], assembled by short (#) and long (*) protocol. Elution was monitored simultaneously by three wavelengths – 280 nm, 350 nm and 415 nm.

CCIS1, the assembly of a [4Fe-4S] cluster significantly increased the α-helix content of the holoprotein with respect to the apoprotein. This is a good example of the power of the computational design algorithm to account for spatial constraints imposed by the [4Fe-4S] cubane on stabilizing the helical structure. In contrast, the empirical design failed to account for the constraints imposed by grafting both types of [4Fe-4S]-binding loops onto a single four-helix bundle although each loop was successfully attached to different four-helix bundles [11,9].

Attempts to improve the computational design of CCIS by intuitive modifications of the protein's sequence that will, for example, increase helix propensity, and reduce charge pairing had very limited success. For most designs, the helix stability of the apoprotein improved, but the main aim, which was to increase the [4Fe-4S] cluster stability over redox cycling, was not reached. The computationally designed sequence of CCIS1 has been optimized by protCAD's computational algorithms. Based on these lessons, we postulate that the design can be improved only by changing the backbone scaffold. We are currently exploring a right-handed coiled coil based on an undecad repeat pattern as an alternative to the left-handed heptad repeat scaffold of CCIS1.

The modularity and simplicity of the empirical approach worked well for improving the design of RCM1. The shortcoming of RCM1 was its high oligomerization tendency. *A posteriori* modeling, however, allowed identifying the asymmetric loop as a problematic point of the structure. Shortening of the loop yielded the RCM2 design which stabilized the apoprotein in a monomeric state, but did not prevent dimerization upon heme binding. This is yet another demonstration of the complexity of designing cofactor-protein complexes. Simple binary patterning of the bundle is not sufficient. Issues like opposing charges on adjacent helices which should work as stabilizers, and the conformational changes induced by ligand binding may require a higher level of computational design with molecular dynamics simulations included.

In conclusion, both applied approaches of protein design have strong and weak points. Both methods enable good predictions of protein structures, but control of ligand behavior is still problematic, as well as controlling the protein's oligomerization state. Computational design is more powerful, being able to reach non-natural folds. Experimental design, however, with a dose of luck and a good physicochemical intuition, may allow creation of systems, which are still out of reach for computational methods, such as mixed cofactors redox chains. Computational methods are developing very fast, and without doubt, in a short time such problems could be also solved. Nonetheless, even very fast and powerful processors need starting points and guidelines, which still rely on examples from natural structural motifs.

Designing new proteins is still a very challenging process, although this field has already featured spectacular successes, such as non-natural folds and enzymes catalyzing reactions not known in nature [20,21].

Acknowledgements

Research was partially supported by Foundation for Polish science, Homing PLUS Programme, co-financed by the European Union within the European Regional Development Fund. JG acknowledge support from EMBO under Short Term Scholarship and NanoFun POIG.02.02.00-00-025/09.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.bbabio.2012.02.001.

References

[1] D.E. Robertson, R.S. Farid, C.C. Moser, J.L. Urbauer, S.E. Mulholland, R. Pidikiti, et al., Design and synthesis of multi-haem proteins, Nature 368 (1994) 425–432.
[2] S. Kamtekar, J.M. Schiffer, H. Xiong, J.M. Babik, M.H. Hecht, Protein design by binary patterning of polar and nonpolar amino acids, Science 262 (1993) 1680–1685.
[3] I. Cohen-Ofri, M. van Gastel, J. Grzyb, A. Brandis, I. Pinkas, W. Lubitz, D. Noy, Zinc-bacteriochlorophyllide dimers in de novo designed four-helix bundle proteins. A model system for natural light energy harvesting and dissipation, J. Am. Chem. Soc. 133 (2011) 9526–9535.
[4] D.E. Benson, M.S. Wisz, H.W. Hellinga, Rational design of nascent metalloenzymes, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 6292–6297.
[5] V. Nanda, M.M. Rosenblatt, A. Osyczka, H. Kono, Z. Getahun, P.L. Dutton, J.G. Saven, W. deGrado, De novo design of a redox-active minimal rubredoxin mimic, J. Am. Chem. Soc. 127 (2005) 5804–5805.
[6] A. Lombardi, F. Nastri, V. Pavone, Peptide-based heme-protein models, Chem. Rev. 101 (2001) 3165–3189.
[7] A.K. Petros, S.E. Shaner, A.L. Costello, D.L. Tierney, B.R. Gibney, Comparison of cysteine and penicillamine ligands in a Co(II) maquette, Inorg. Chem. 43 (2004) 4793–4795.
[8] M.L. Kennedy, A.K. Petros, B.R. Gibney, Cobalt(II) and zinc(II) binding to a ferredoxin maquette, J. Inorg. Biochem. 98 (2004) 727–732.
[9] M.P. Scott, J. Biggins, Introduction of a [4Fe-4S (S-cys)4]+1,+2 iron-sulfur center into a four-alpha helix protein using design parameters from the domain of the Fx cluster in the Photosystem I reaction center, Protein Sci. 6 (1997) 340–346.
[10] M.L. Kennedy, B.R. Gibney, Proton coupling to [4Fe-4S](2+/+) and [4Fe-4Se](2+/+) oxidation and reduction in a designed protein, J. Am. Chem. Soc. 124 (2002) 6826–6827.
[11] B.R. Gibney, S.E. Mulholland, F. Rabanal, P.L. Dutton, Ferredoxin and ferredoxin-heme maquettes, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 15041–15046.
[12] S.M. Lippow, B. Tidor, Progress in computational protein design, Curr. Opin. Biotechnol. 18 (2007) 305–311.

[13] A.D. Keefe, J.W. Szostak, Functional proteins from a random-sequence library, Nature 410 (2001) 715–718.
[14] J. Grzyb, F. Xu, L. Weiner, E.J. Reijerse, W. Lubitz, V. Nanda, D. Noy, De novo design of a non-natural fold for an iron-sulfur protein: alpha-helical coiled-coil with a four-iron four-sulfur cluster binding site in its central core, Biochim. Biophys. Acta 1797 (2010) 406–413.
[15] Y. Wei, M.H. Hecht, Enzyme-like proteins from an unselected library of designed amino acid sequences, Protein Eng. Des. Sel. 17 (2004) 67–75.
[16] T. Kortemme, D. Baker, Computational design of protein–protein interactions, Curr. Opin. Chem. Biol. 8 (2004) 91–97.
[17] B. Kuhlman, G. Dantas, G.C. Ireton, G. Varani, B.L. Stoddard, D. Baker, Design of a novel globular protein fold with atomic-level accuracy, Science 302 (2003) 1364–1368.
[18] G. Ghirlanda, J.D. Lear, A. Lombardi, W.F. DeGrado, From synthetic coiled coils to functional proteins: automated design of a receptor for the calmodulin-binding domain of calcineurin, J. Mol. Biol. 281 (1998) 379–391.
[19] M.A. Dwyer, L.L. Looger, H.W. Hellinga, Computational design of a biologically active enzyme, Science 304 (2004) 1967–1971.
[20] D. Rothlisberger, O. Khersonsky, A.M. Wollacott, L. Jiang, J. DeChancie, J. Betker, et al., Kemp elimination catalysts by computational enzyme design, Nature 453 (2008) 190–195.
[21] L. Jiang, E.A. Althoff, F.R. Clemente, L. Doyle, D. Rothlisberger, A. Zanghellini, J.L. Gallaher, J.L. Betker, F. Tanaka, C.F. Barbas, D. Hilvert, K.N. Houk, B.L. Stoddard, D. Baker, De novo computational design of retro-aldol enzymes, Science 319 (2008) 1387–1391.
[22] L. Yuan, I. Kurek, J. English, R. Keenan, Laboratory-Directed Protein Evolution, Microbiol. Mol. Biol. Rev. 69 (2005) 373–392.
[23] J.R. Beasley, M.H. Hecht, Protein design: the choice of de novo sequences, J. Biol. Chem. 272 (1997) 2031–2034.
[24] G. Shen, J.H. Golbeck, Assembly of the bound iron-sulfur clusters in photosystem I, Photosystem i - The Light-Driven Plastocyanin: Ferredoxin Oxidoreductase, 24, 2006, pp. 529–547.
[25] T. Friedrich, K. Steinmuller, H. Weiss, The proton-pumping respiratory complex I of bacteria and mitochondria and its homologue in chloroplasts, FEBS Lett. 367 (1995) 107–111.
[26] M.L. Ghirardi, M.C. Posewitz, P.C. Maness, A. Dubini, J. Yu, M. Seibert, Hydrogenases and hydrogen photoproduction in oxygenic photosynthetic organisms, Annu. Rev. Plant Biol. 58 (2007) 71–91.
[27] M. Ihara, H. Nishihara, K.S. Yoon, O. Lenz, B. Friedrich, H. Nakamoto, K. Kojima, D. Honma, T. Kamachi, I. Okura, Light-driven hydrogen production by a hybrid complex of a [NiFe]-hydrogenase and the cyanobacterial photosystem I, Photochem. Photobiol. 82 (2006) 676–682.
[28] A.J. Doig, R.L. Baldwin, N- and C-capping preferences for all 20 amino acids in alpha-helical peptides, Protein Sci. 4 (1995) 1325–1336.
[29] M.J.I. Andrews, A.B. Tabor, Forming stable helical peptides using natural and artificial amino acids, Tetrahedron 55 (1999) 11711–11743.
[30] G. Ghirlanda, A. Osyczka, W. Liu, M. Antolovich, K.M. Smith, P.L. Dutton, A.J. Wand, W.F. deGrado, De novo design of a d2-symmetrical protein that reproduces the diheme four-helix bundle in cytochrome bc1, J. Am. Chem. Soc. 126 (2004) 8141–8147.
[31] C.D. Stout, E.A. Stura, D.E. McRee, Structure of Azotobacter vinelandii 7Fe ferredoxin at 1.35 A resolution and determination of the [Fe-S] bonds with 0.01 A accuracy, J. Mol. Biol. 278 (1998) 629–639.
[32] A. Amunts, O. Drory, N. Nelson, The structure of a plant photosystem I supercomplex at 3.4 A resolution, Nature 447 (2007) 58–63.
[33] A. Roy, A. Kucukural, Y. Zhang, I-TASSER: a unified platform for automated protein structure and function prediction, Nat. Protoc. 5 (2010) 725–738.
[34] C.M. Summa, Computational methods and their applications for de novo functional protein design and membrane protein solubilization, Doctoral Thesis, University of Pennsylvania School of Medicine, Philadelpha, (2002).
[35] M.L. Antonkine, E.M. Maes, R.S. Czernuszewicz, C. Breitenstein, E. Bill, C.J. Falzone, R. Balasubramanian, C. Lubner, D.A. Bryant, J.H. Golbeck, Chemical rescue of a site-modified ligand to a [4Fe-4S] cluster in PsaC, a bacterial-like dicluster ferredoxin bound to Photosystem I, Biochim. Biophys. Acta 1767 (2007) 712–724.
[36] J.D. Weinstein, R.W. Howell, R.D. Leverette, S.Y. Grooms, P.S. Brignola, S.M. Mayer, S.I. Baele, Heme Inhibition of [delta]-Aminolevulinic Acid Synthesis Is Enhanced by Glutathione in Cell-Free Extracts of Chlorella, Plant Physiol. 101 (1993) 657–665.
[37] R.W. Sreerama, N. Woody, Computation and analysis of protein circular dichroism spectra. Methods Enzymol. 383 (2004) 318–351.
[38] P.W. Riddles, R.L. Blakeley, B. Zerner, Reassessment of Ellman's reagent, in: S.N.T.C.H.W. Hirs (Ed.), Methods in Enzymology, Academic Press, 1983, pp. 49–60.
[39] W.V. Sweeney, J.C. Rabinowitz, Proteins containing 4Fe-4S clusters: an overview, Annu. Rev. Biochem. 49 (1980) 139–161.